



<i>Acknowledgement</i>	This is an Accepted Manuscript of an article published by Taylor & Francis in <i>Journal of Child Sexual Abuse</i> on 12 May 2014, available online.
<i>Archived version</i>	Author manuscript: post-print (post-refereeing)
<i>Published version</i>	http://dx.doi.org/10.1080/10538712.2014.896840
<i>Journal homepage (Copyright)</i>	http://www.tandfonline.com/toc/wcsa20/current ISSN: 1053-8712, EISSN: 1547-0679
<i>Author contact at KU Leuven</i>	kai.karos@ppw.kuleuven.be

(article begins on next page)

Factor Structure, Reliability, and Known Groups Validity of the German Version of the
Childhood Trauma Questionnaire (Short-Form) in Swiss Patients and Non-Patients

Kai Karos, Nils Niederstrasser, Latifa Abidi and David P. Bernstein

Maastricht University, Maastricht, The Netherlands

Klaus Bader

Universität Basel, Basel, Switzerland

Author Note

Kai Karos, Department of Psychology and Neuroscience, Maastricht University; Nils Niederstrasser, Department of Psychology and Neuroscience, Maastricht University; Latifa Abidi, Department of Psychology and Neuroscience, Maastricht University; David P. Bernstein, Department of Clinical Psychological Science, Maastricht University; Klaus Bader, Department for Clinical Psychology and Psychotherapy, Basel University.

Kai Karos is now at Research Group on Health Psychology, University of Leuven; Nils Niederstrasser is now at Department of Psychology, Research Group on Health Psychology, University of Leuven; Latifa Abidi is now at Caphri School for Public Health and Primary Care, Department of Health Services Research, Maastricht University; David P. Bernstein, Department of Clinical Psychological Science, Maastricht University; Klaus Bader, Department for Clinical Psychology and Psychotherapy, Basel University.

Correspondence concerning this article should be addressed to Kai Karos, Research Group on Health Psychology, University of Leuven, P.O. Box 3726, 3000 Leuven, Belgium.
E-mail: kai.karos@ppw.kuleuven.be

Abstract

The Childhood Trauma Questionnaire-Short Form (CTQ-SF) is the most widely used instrument to assess childhood trauma and has been translated into 10 languages. However, research into validity and reliability of these translated versions is scarce. The present study aimed to investigate the factor structure, internal consistency, reliability and known-groups validity of the German CTQ-SF (Bernstein & Fink, 1998). Six-hundred sixty-one clinical and non-clinical participants completed the German CTQ-SF. A confirmatory factor analysis (CFA) was conducted to assess the 5-factor structure of the original CTQ-SF. To investigate known-groups validity, the CFA latent factor levels between clinical and non-clinical participants were compared. The original 5-factor structure was confirmed, with only the Physical Neglect scale showing rather poor fit. As a conclusion, the results support the validity and reliability of the German CTQ-SF. It is recommended to use the German CTQ-SF to assess experiences of childhood trauma.

Keywords: children, sexual abuse, methodology, psychometric, validity, reliability, CTQ-SF

Research in recent years suggests that psychopathology seems to be the result of complex interactions between environmental risk factors and genetic vulnerability (Caspi & Moffitt, 2006; Kim-Cohen et al., 2006). One of the most prominent and frequently proposed environmental risk factors is the presence of traumatic experience during childhood, especially child abuse and neglect. Unfortunately, the prevalence of traumatic experiences in childhood is alarming. Across different community samples, experience of at least one form of abuse in childhood was 26.6% for females and 31.7% for males (MacMillan et al., 1997). In the general population, prevalence for sexual abuse was 14.2% for men and 32.3% for women, and physical abuse 19.5% for women and 22.2% for men (Briere & Elliott, 2003; MacMillan et al., 2001). Research has found that the consequences of childhood maltreatment are numerous and serious, such as higher risk for the development of several psychopathologies such as depression, substance abuse, post-traumatic stress disorder (PTSD) and personality disorders. Other possible consequences of childhood maltreatment include aggression and antisocial behavior, sexually inappropriate behavior, anxiety, self-destructive behavior, and revictimization (Briere & Elliott, 2003; Browne & Finkelhor, 1986; Mulvihill, 2005; Putnam, 2003). Consequently, scientific and clinical practice is in great need of valid and reliable measures that are able to assess traumatic experiences, such as abuse and neglect, in childhood.

The most prominent instrument to measure the experience of childhood trauma is the Childhood Trauma Questionnaire-Short Form (CTQ-SF; Bernstein & Fink, 1998). The CTQ-SF is a retrospective 28-item self-report questionnaire that measures the exposure to traumatic experiences in childhood. It assesses childhood maltreatment on five distinct dimensions: Physical Abuse, Emotional Abuse, Sexual Abuse, Physical Neglect, and Emotional Neglect. The CTQ-SF has been shown to have good reliability and validity in clinical as well as community samples (Bernstein et al., 2003; Scher, Stein, Asmundson, McCreary, & Forde, 2001). The 5-factor structure of the CTQ-SF, has been confirmed in

several studies (Bernstein et al., 2003; Scher et al., 2001). Furthermore, the CTQ-SF has been shown to have good convergent validity with other measures that assess maltreatment retrospectively. Moreover, good criterion validity with therapists' ratings of abuse and neglect in adolescents, many of whom had confirmed abuse and neglect histories, has been shown (Bernstein et al., 2003).

The CTQ-SF has been translated into many different languages, among them Dutch (Arntz, 1999), Spanish (Basurte, Diaz-Marsa, Martin, & Carrasco, 2004), Turkish (Aslan & Alparslan, 1999) and German (Bader, Schäfer, Schenkel, Nissen, & Schwandler, 2007). Although the original English CTQ-SF is the most validated questionnaire of its kind, there have been few attempts to investigate the validity and reliability of the CTQ-SF in different language, cultural, or gender groups. A recent study by Thombs, Bernstein, Lobbestael, and Arntz (2009) investigated the structural validity of the translated Dutch CTQ-SF, namely testing factor structure, internal consistency reliability, and known-groups validity. It has been shown that the translated version of the CTQ-SF has the same 5-factor structure as the original English version of the questionnaire. However, a single item turned out to be mistranslated, which was shown by aberrant factor loadings. In general, there are a number of very important reasons why a validation study of a translated instrument is worthwhile. First, even if back translation was used, one cannot assume that all items retain their original meaning, as understood by respondents across language groups. Statistical techniques like confirmatory factor analysis are therefore employed to detect deviations between the original and the translated scale, as shown in the study by Thombs et al. (2009).

Furthermore, culture as well as gender issues might influence the interpretation of items and the reporting of abuse and neglect. For instance, what is considered maltreatment might differ substantially between cultures. For example, in one study, African-Americans were less likely to define certain forms of physical maltreatment as abuse, and men were less likely to subjectively define themselves as physically abused, even when objective items

indicated that they had been (Thombs et al., 2009). Consequently it is crucial to investigate whether the interpretation of items differs across cultures.

Scales that are used worldwide also serve the function to enable comparisons between different cultures. For example, comparisons of the prevalence of different forms of maltreatment. In order to make meaningful multicultural comparisons, measurement invariance has to be established (Vandenberg & Lance, 2000). Measurement invariance refers to the equivalence of a scale's factor structure across different populations and is a precondition to use an instrument to compare different samples. Only when the translated instruments have been tested in terms of validity and reliability, can they be meaningfully used for measurement; thus, assessing the same underlying construct as the original instrument.

The aim of this study is to assess the reliability and validity of the translated German version of the CTQ-SF in a Swiss-German sample. The study assesses the factor structure of the CTQ-SF with confirmatory factor analysis, and investigates the known-groups validity by comparing the scores on the CTQ-SF of a clinical to a non-clinical sample. Although Switzerland is divided into 26 provinces and four official languages are spoken, only participants whose first language is German were recruited for this study. Although spoken Swiss-German differs from German as it is spoken in Germany, the written German is the same for both countries. Consequently, the results of this study should be applicable not only for Switzerland, but for all countries where written German is used officially. In a previous study, Bader, Hänni, Schäfer, Neuckel, and Kuhl (2009) reported that a German translation of the CTQ-SF had the same 5-factor structure as the original English language version, although the reliability of one of the factors, Physical Neglect, was poor. However, the sample size used in that study was smaller ($N = 363$). The present study is mainly an extension of this earlier study. The main extension will be the use of a larger sample and the inclusion of a non-clinical population. That way it will be possible to compare the CTQ-SF

across clinical and non-clinical populations (“known-groups validity”). Furthermore, the conclusions of the present study can be extended to a mixed sample rather than to a clinical sample alone. As German is used by 90 to 98 million native speakers worldwide, this study aims to determine whether the CTQ-SF is reliable and valid for this large population.

Method

Sample Selection and Procedure

A total of 661 individuals participated in this study: 565 with clinical disorders and 96 Psychology students. The clinical sample consisted of three different clinical populations. The first and largest sample was taken from the behavioral therapy university clinic in Basel ($n = 487$). Patients from this clinic mainly suffer from anxiety disorders or depression. The CTQ-SF was part of their standard diagnostic procedures. The second sample was recruited from a forensic university clinic in Basel ($n = 18$). These patients were mainly diagnosed with pedophilia. The staff of the forensic institution asked these patients whether they would be interested to fill in the CTQ-SF as part of a study that includes the completion of various questionnaires. The third sample comprised of individuals diagnosed with sleep disorders (mainly insomnia) recruited from the general population ($n = 60$). Participants for the latter study were recruited via newspaper advertisements. The study investigated the relationship between sleep disorders and traumatic experiences in childhood (Bader et al., 2007). Of the 661 participants in our study, 424 (363 patients and 61 students) had participated in the earlier reported study of Bader and colleagues (2009).

The studies for samples two and three were approved by the ethical committee of the University of Basel, whereas this was not necessary for the patients of sample one, because the CTQ-SF was used as part of the routine clinical assessment of patients. The healthy controls were recruited from the University of Basel. All non-patients were clinical psychology students enrolled in a master’s program (see Table 1). They completed the CTQ-

SF as part of an obligatory seminar in their master's program. All participants signed an informed consent.

Table 1

Demographic and Clinical Characteristics for Clinical and Non-Clinical Respondents

	<u>Total sample</u> (<i>N</i> = 661)	<u>Clinical sample</u> (<i>n</i> = 565)	<u>Non-clinical sample</u> (<i>n</i> = 96)	<i>p</i> -value
Age in years, <i>mean (SD)</i>	35.08(11.747)	36.64(11.79)	25.70(5.43)	< .001
Highest Education Attained				> .05 ¹
Compulsory school, <i>n (%)</i>	8 (1.2)	8(1.4)	0 (.0)	
Junior high school, <i>n (%)</i>	108(16.3)	108(19.1)	0 (.0)	
Apprenticeship, <i>n (%)</i>	253(38.3)	253(44.8)	0 (.0)	
High school, <i>n (%)</i>	174(26.3)	81(14.3)	93(96.9)	
Completed University/college, <i>n (%)</i>	110(16.6)	107(18.9)	3(3.1)	
Occupation				< .001 ²
In training, <i>n (%)</i>	179(27.1)	96(17)	83(86.5)	
Household, <i>n (%)</i>	21(3.2)	21(3.7)	0 (.0)	
Part-time, <i>n (%)</i>	143(21.6)	130(23.0)	13(13.5)	
Fulltime, <i>n (%)</i>	167(25.3)	167(29.6)	0 (.0)	
Retired, <i>n (%)</i>	8(1.2)	8(1.4)	0 (.0)	
Unemployed, <i>n (%)</i>	47(7.1)	47(8.3)	0 (.0)	
Sick, <i>n (%)</i>	16(2.4)	16(2.8)	0 (.0)	
Invalid pension, <i>n (%)</i>	62(9.4)	62(11.0)	0 (.0)	
Social welfare, <i>n (%)</i>	14(2.1)	14(2.5)	0 (.0)	
Other, <i>n (%)</i>	1(0.2)	1(0.2)	0 (.0)	
Therapy status				
Outpatient, <i>n (%)</i>	554(83.8)	554(98.1)	0 (.0)	
Inpatient, <i>n (%)</i>	5(0.8)	5(0.9)	0 (.0)	
In- and outpatient, <i>n (%)</i>	6(0.9)	6(11.1)	0 (.0)	
CTQ-SF scale scores				
Physical abuse, <i>mean(SD)</i>	6.01(2.58)	6.13(2.73)	5.3(1.22)	.003
Emotional abuse, <i>mean(SD)</i>	8.01(3.87)	8.25(4.04)	6.54(2.22)	< .001
Sexual abuse, <i>mean(SD)</i>	6.26(3.45)	6.41(3.63)	5.39(1.82)	.007
Physical neglect, <i>mean(SD)</i>	7.28(2.58)	7.53(2.64)	5.84(1.61)	< .001
Emotional neglect, <i>mean(SD)</i>	12.99(5.64)	13.58(5.66)	9.49(4.07)	< .001

The CTQ-SF

The 28 items of the CTQ-SF assess five types of childhood trauma (physical and emotional abuse, physical and emotional neglect and sexual abuse) with 5 items for each scale and 3 items for a minimization/denial validity scale. The possible responses on each item range from 1 to 5 and reflect the frequency of maltreatment experiences with 1 being

Note. ¹High-school (or higher) per group [$\chi^2(1) = 1.4$, $p = .237$]; ²Working vs. Not working per group [$\chi^2(1) = 30.92$, $p = .000$]

never true and 5 being *very often true*. For the content of the items see Table 1. This 5-factor structure has been confirmed in multiple studies with respect to the original CTQ-SF (Bernstein et al., 2003). In addition to the evidence for the validity of the CTQ-SF noted above, earlier studies found generally good reliability ($\alpha > .80$) for all of the CTQ-SF's scales with the exception of the Physical Neglect scale, which typically has reliability in the range from .60 to .70 (Bernstein et al., 2003; Scher et al., 2001; Thombs et al., 2009).

Back translation. Back translation was independently performed by two masters students in psychology (K.K. and N.N.), both of whom are native German speakers, fluent in English, and blind to the original version of the CTQ-SF. The two back-translated versions were inspected by David P. Bernstein, the author of the CTQ-SF, and compared to the English language original. There were only slight discrepancies in a few of the items between the original and back-translated versions, none of which altered the meaning of the original items. Thus, the German translation of the CTQ-SF appeared to be consistent with the original.

Data Analysis

The confirmatory factor analysis (CFA) was done with LISREL testing the validity of the 5-factor structure of the original CTQ-SF and assessing the loading of the distinct items on each of the different subscales. To assess model fit, a chi-square goodness-of-fit test was used along with three other fit indices, namely the CFI, the SRMR and the RMSEA. According to the guidelines by Hu and Bentler (1999), a CFI close to .95 or higher, a SRMR close to .08 or lower, and a RMSEA between .06 to .08 or lower represent a good model fit and are acceptable. All items responses were ordinal Likert data, and in order to correct for skewedness in the item responses, square root transformation was used. All statistical analyses were done with the transformed item responses.

Comparisons for demographic and clinical characteristics between clinical and non-clinical respondents were assessed using independent samples *t*-tests and the Chi-square

statistic. To assess internal consistency reliability, Cronbach's Alpha was computed for each CTQ-SF scale. For the latter analyses, SPSS-15 was used. To assess known-groups validity, the latent factor levels in the CFA model for each scale were compared between patients and non-patients.

Results

Sample characteristics

Demographic and clinical characteristics of the study participants are shown in Table 1. It was found that the mean age in the clinical sample was significantly higher than in the non-clinical sample, $t(657) = -8.8, p < .001$. Moreover, the clinical sample scored higher on all given scales of the CTQ-SF compared to the non-clinical sample. That is, they scored higher on Physical Abuse $t(659) = -2.9, p = .003$; on Emotional Abuse $t(659) = -4.1, p < .001$; on Sexual Abuse $t(659) = -2.7, p = .007$; on Emotional Neglect $t(659) = -6.8, p < .001$; and on Physical Neglect $t(659) = -6, p < .001$.

CTQ-SF item responses and factor structure

The standard five-factor model based on the original English CTQ was tested initially. Based on this analysis, three item error covariances were freed based on modification indices. All members of each pair were from the same scale and shared method and format features. In line with the original version of the CTQ-SF (Thombs et al., 2007), and the Dutch translation of the CTQ-SF (Thombs et al., 2009), error variances were freed to covary for items PN1 and PN6, EA3 and EA14 and EN13 and EN19 (see Table 2). This resulted in a well-fitting model [$\chi^2(262) = 884.12$, CFI = .97, SRMR = .06, RMSEA = .06] and in overall high factor loadings. The only exception are items PN1, PN4 and PN6 with all having loadings $< .3$ (see Table 2). Although factor loadings on the Physical Neglect scale have been found to be the lowest of the CTQ before (Thombs et al., 2009), the loadings in this case

are exceptionally low. The correlations between the different kinds of abuse were moderate, except the correlation between Emotional and Physical Neglect (.89; see Table 3).

Table 2

Standardized Factor Loadings for 25-item German Childhood Trauma Questionnaire-Short Form

	Physical Abuse	Emotional Abuse	Sexual Abuse	Physical Neglect	Emotional Neglect
Physical Abuse					
PA9-Hit hard enough to see doctor	.65				
PA11-Hit hard enough to leave bruises	.87				
PA12-Punished with hard objects	.74				
PA15-Was physically abused	.71				
PA17-Hit badly enough to be noticed	.52				
Emotional Abuse					
EA3-Called names by family		.65			
EA8-Parents wished was never born		.68			
EA14-Family said hurtful things		.75			
EA18-Felt hated by family		.71			
EA25-Was emotionally abused		.65			
Sexual Abuse					
SA20-Was touched sexually			.89		
SA21-Hurt if did not do something sexual			.67		
SA23-Made to do sexual things			.75		
SA24-Being molested			.88		
SA27-Was sexually abused			.84		
Physical Neglect					
PN1-Not enough to eat				.24	
PN2(R)-Got taken care of				.86	
PN4-Parents were high or drunk				.26	
PN6-Wore dirty clothes				.20	
PN26(R)-Got taken to doctor				.56	
Emotional Neglect					
EN5(R)-Made to feel important					.68
EN7(R)-Felt loved					.85
EN13(R)-Was looked out for					.84
EN19(R)-Family felt close					.85
EN28(R)-Family was source of strength					.90

Note. (R) indicates reverse coded and scored.

Table 3

Correlations between Factors

	Emotional Abuse	Physical Abuse	Sexual Abuse	Emotional neglect	Physical neglect
Emotional Abuse	1.00				
Physical Abuse	0.68	1.00			
Sexual Abuse	0.33	0.33	1.00		
Emotional neglect	0.64	0.38	0.25	1.00	
Physical neglect					1.00

Physical neglect	0.57	0.42	0.27	0.89	1.00
------------------	------	------	------	------	------

Reliability

The internal consistency was assessed using Cronbach's Alpha for each CTQ-SF scale. Cronbach's Alpha was .82 for Physical Abuse, .83 for Emotional Abuse, .90 for Sexual Abuse, .91 for Emotional Neglect, and .53 for Physical Neglect. Thus, the Physical Neglect scale seems to have low internal consistency reliability.

Known-groups validity

In order to test the known-groups validity of the German version of the CTQ-SF, the final CFA model was rerun with each factor regressed on clinical status, that is, clinical versus non-clinical groups (see Thombs et al., 2009). The model fit decreased [$\chi^2(292) = 1995.84$, CFI = .92, SRMR = .2, RMSEA = .09]. Still, factor loadings were higher for patients compared to controls: Physical Abuse (0.33 SD), Emotional Abuse (0.49 SD), Sexual Abuse (0.33 SD), Physical Neglect (0.82 SD) and Emotional Neglect (0.73 SD).

Discussion

Based on the present study, several main conclusions can be drawn. First, this study found that the original five-factor model of the CTQ-SF applies to the German translation of the CTQ-SF. Second, the CTQ-SF German version revealed satisfactory internal consistency reliability among its scales with the exception of the Physical Neglect scale, which was lower than in most published studies of the CTQ-SF (Bernstein et al., 2003). Additionally, this scale showed the weakest factor loadings. Prior studies have revealed that this particular scale is commonly found to be the weakest factor of the five-factor solution in terms of reliability and factor loadings (Thombs, et al., 2009). Also, this scale shows a high correlation with the Emotional Neglect scale, even higher than in other published studies. One possible explanation would be that this is a consequence of the translation itself. However, this seems unlikely, given that the back-translation was adequate. Moreover, the

items loaded on the intended factors, unlike in Thombs and colleagues (2009) analysis of the Dutch translation where one of the items, which had been mistranslated, loaded highly on two factors. An alternative explanation is that the base rates of physical neglect were very low in this sample. Switzerland has one of the highest standards of living in the world, as well as a robust social welfare safety net for those who are in need. This may explain why very few individuals in this sample endorse items involving not having enough food to eat or having to wear dirty clothes. In this sense, physical neglect may be a difficult concept to differentiate from poverty and its consequences.

Moreover, the known-groups validity analysis has produced somewhat mixed results. Although the model fit decreased if the distinction between known groups is entered in the model, as indicated by the CFI (.92) and the RMSEA (.09), the model fit can still be considered indicative of group differences. A possible reason why the known-groups validity is less pronounced in the present sample might be a smaller difference with regard to trauma experiences in the two groups (about one third to three quarters of a standard deviation difference compared to the sample analysed by Thombs et al., 2009). In conclusion, the present findings provide some support for known groups validity, but not as unequivocally as in earlier validation studies of the CTQ-SF. As a comparison, latent factor levels were more than one standard deviation higher for patients compared to controls in the study by Thombs et al. (2009).

Several limitations should be considered when interpreting the present results. As with nearly all retrospective studies of child abuse and neglect, there was no means to independently verify the self-report responses on the CTQ-SF. Comparing participants' responses on the CTQ-SF to corroborative information, such as child welfare records, or reports on other trauma measures, such as structured trauma interviews, would have been desirable. Secondly, the use of a mainly clinical sample is a limitation of the present study. In order to strengthen the generalizability of the German CTQ-SF, it would have been

desirable to have equally large samples of both clinical as well as healthy participants. Most conservatively, one could conclude that the findings presented in this study support the factorial validity of the CTQ in a mixed sample, mainly consisting of clinical patients. In a future study a larger sample of healthy participants would allow for a direct comparison of the factor structure in clinical and non-clinical individuals. Thirdly, it would have been desirable to include a German speaking non-Swiss sample in the study, to test if the results can be generalized to non-Swiss German speakers. Although on theoretical grounds a difference between these samples is unlikely, as the written German language is the same in both countries, it would be worthwhile to support this with empirical evidence. Also, ideally, one should compare German and English language samples on the CTQ-SF in the same study. Doing so would enable one to directly compare the size of the factor loadings to investigate possible differences between the two language groups. However, it is difficult to find comparable samples in different language groups that are very similar in terms of their demographic or clinical characteristics.

To summarize, this study shows good fit of the five-factor model of the German language version of the CTQ-SF and good reliability with the exception of Physical Neglect scale, which was fair. In general, however, this study suggests that the psychometric properties of the German CTQ-SF are comparable to the English original, making it suitable for both clinical and research assessment of child abuse and neglect in German speaking patients and non-patients.

References

- Arntz, A. (1999). Do personality disorders exist? On the validity of the concept and its cognitive behavioral formulation and treatment. *Behaviour Research and Therapy*, 37, 97–134. doi:10.1016/S0005-7967(99)00052-2
- Aslan, S. H., & Alparslan, Z. N. (1999). The reliability, validity and factor structure of the Childhood Trauma Questionnaire among a group of university students. *Turkish Journal of Psychiatry*, 10(4), 275-285. Retrieved from <http://www.turkpsikiyatri.com/en/default.aspx?modul=summary&id=110>
- Bader, K., Schäfer, V., Schenkel, M., Nissen, L., & Schwandler, J. (2007). Adverse childhood experiences associated with sleep in primary insomnia. *Journal of Sleep Research*, 16(3), 285–296.
- Bader, K., Hanny, C., Schäfer, V., Neuckel, A., & Kuhl, C. (2009). Childhood Trauma Questionnaire – Psychometric properties of a German version. *Journal for Clinical Psychology and Psychotherapy*, 38(4), 223–230. doi:10.1026/1616-3443.38.4.223
- Basurte, E., Diaz-Marsa, M., Martin, O., & Carrasco, J. L. (2004). [Traumatic childhood background, impulsiveness and hypothalamus-pituitary-adrenal axis dysfunction in eating disorders. A pilot study]. *Spanish Journal of Psychiatry Actas espanolas de psiquiatria*, 32(3), 149-152. Retrieved from <http://europepmc.org/abstract/MED/15168265>
- Bernstein, D. P., & Fink, L. (1998). *Childhood Trauma Questionnaire: A retrospective self-report*. San Antonio, TX: The Psychological Corporation.
- Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T... Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse & Neglect*, 27(2), 169–190.

- Briere, J., & Elliott, D. M. (2003). Prevalence and psychological sequelae of self-reported childhood physical and sexual abuse in a general population sample of men and women. *Child Abuse & Neglect*, 27(10), 1205–1222.
- Browne, A., & Finkelhor, D. (1986). Impact of child sexual abuse: A review of the research. *Psychological Bulletin*, 99(1), 66.
- Caspi, A., & Moffitt, T. E. (2006). Gene–environment interactions in psychiatry: Joining forces with neuroscience. *Nature Reviews Neuroscience*, 7(7), 583–590.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- Kim-Cohen, J., Caspi, A., Taylor, A., Williams, B., Newcombe, R., Craig, I. W., & Moffitt, T. E. (2006). MAOA, maltreatment, and gene–environment interaction predicting children’s mental health: New evidence and a meta-analysis. *Molecular psychiatry*, 11(10), 903–913.
- MacMillan, H. L., Fleming, J. E., Streiner, D. L., Lin, E., Boyle, M. H., Jamieson, E.,... Beardslee, W. R. (2001). Childhood abuse and lifetime psychopathology in a community sample. *American Journal of Psychiatry*, 158(11), 1878–1883.
- MacMillan, H. L., Fleming, J. E., Trocmé, N., Boyle, M. H., Wong, M., Racine, Y. A.,... Offord, D. R. (1997). Prevalence of child physical and sexual abuse in the community. *JAMA: The Journal of the American Medical Association*, 278(2), 131–135.
- Mulvihill, D. (2005). The health impact of childhood trauma: An interdisciplinary review, 1997-2003. *Issues in Comprehensive Pediatric Nursing*, 28(2), 115–136.
- Putnam, F. W. (2003). Ten-year research update review: Child sexual abuse. *Journal of the American Academy of Child & Adolescent Psychiatry*, 42(3), 269–278.

Scher, C. D., Stein, M. B., Asmundson, G. J. G., McCreary, D. R., & Forde, D. R. (2001).

The childhood trauma questionnaire in a community sample: Psychometric properties and normative data. *Journal of Traumatic Stress, 14*(4), 843–857.

Thombs, B. D., Bernstein, D. P., Lobbestael, J., & Arntz, A. (2009). A validation study of the Dutch Childhood Trauma Questionnaire-Short Form: Factor structure, reliability, and known-groups validity. *Child Abuse & Neglect, 33*(8), 518–523.

doi:10.1016/j.chiabu.2009.03.001

Thombs, B. D., Lewis, C., Bernstein, D. P., Medrano, M. A., & Hatch, J. P. (2007). An evaluation of the measurement equivalence of the Childhood Trauma Questionnaire-Short Form across gender and race in a sample of drug-abusing adults. *Journal of Psychosomatic Research, 63*(4), 391–398. doi:10.1016/j.jpsychores.2007.04.010

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4–70.